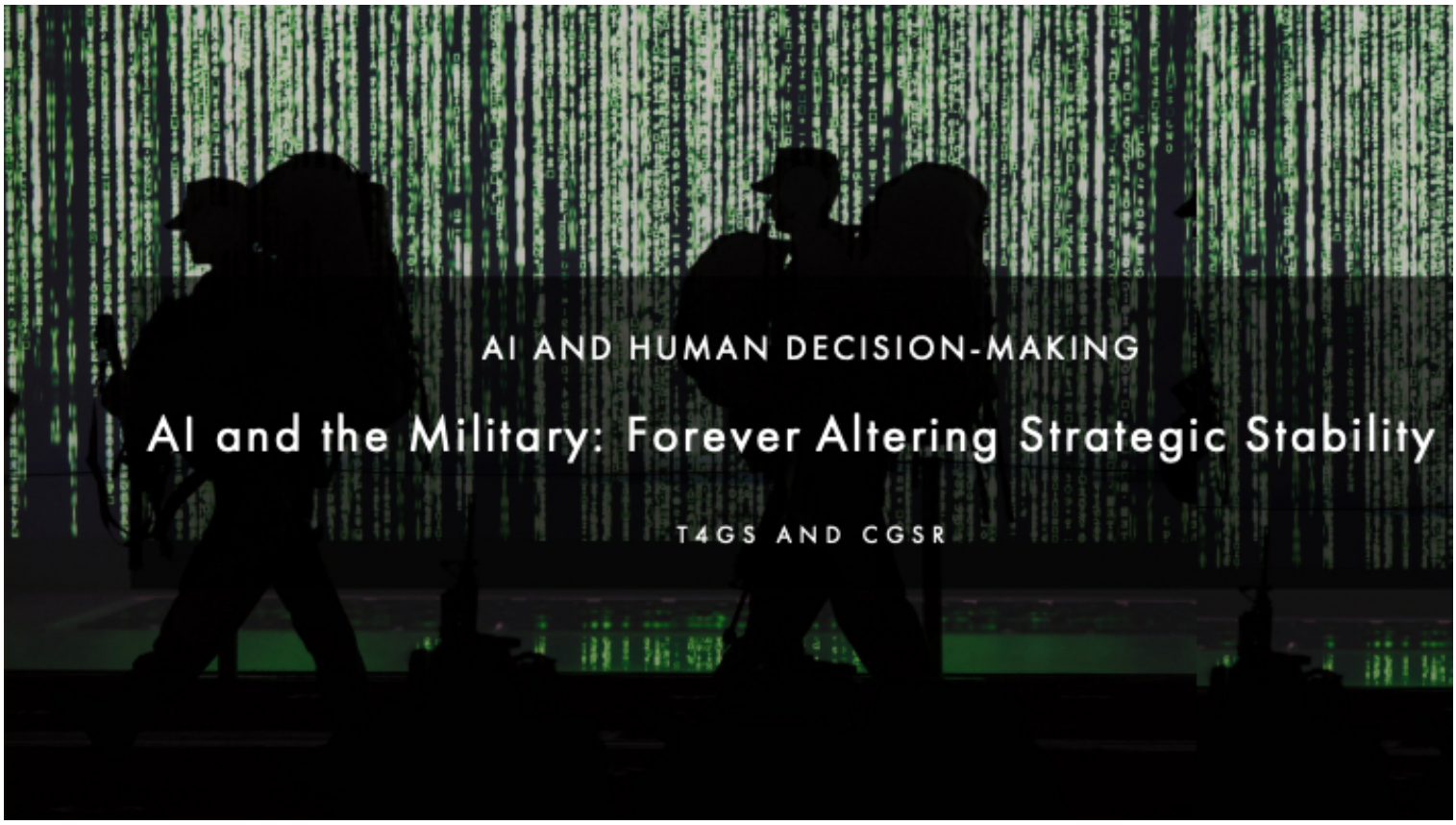
TECHNOLOGY FOR GLOBAL SECURITY

AI and the Military: Forever Altering Strategic Stability

Technology for Global Security and the Center for Global Security Research

February 13, 2019



Recommended Citation:

T4GS, "AI and the Military: Forever Altering Strategic Stability", T4GS Reports, February 13, 2019, http://www.tech4gs.org/ai-and-human- decision-making.html

AI and the Military: Forever Altering Strategic Stability

Artificial intelligence has burst upon the national security scene with an intensity to surprise even the most veteran observers of the national policy discourse. The renewed spike of interest is driven in part by popular characterizations of novel AI techniques as revolutionary, ostensibly on par with the discovery of fire, electricity, or nuclear weapons.[1] It is also driven in part by the rapid absorption of nascent AI-based technologies - primarily driven by novel machine learning techniques - into diverse sectors of the global economy, often with transformative effects (as for example in the sciences and in social media). It is also driven, however, in large part by the great power ambitions of America's competitors and potential adversaries. Echoing the 19th century naval strategist Alfred Mahan (*"Whoever rules the waves rules the world"),* Russia's President Putin has stated that *the nation that rules in AI "will be the ruler of the world."*[2] China's leader Xi Jinping is perhaps less demonstrative publicly, but has committed to making China the dominant global AI power by 2030, following what was widely heralded as China's "Sputnik moment" after the loss by Lee Sedol to AlphaGo in 2016.[3] There are mounting fears that the United States is woefully under-prepared to manage these new challenges, and that the United States will end up "offset" due to the sheer scale at which the Chinese intend to deploy AI. Could AI disrupt and reshape the strategic international balance, as blue water navies and nuclear weapons did in previous eras? Might it do so in a manner so severe that it leads to war? Is an AI arms race underway?[4] Will imbalances and changing perceptions of capabilities undermine the status quo of what is needed to maintain strategic stability between near-peer powers?

The purpose of this paper is to begin to answer some of these questions and contribute to the growing body of research and analysis, while calibrating the potential risks and rewards of military applications of AI technologies and determining which issues demand further research - and action. To do so, this paper explores the following questions:

1. Which technologies have potential near-term military applications, and which do not?
2. Of those, which are potentially consequential for strategic stability? How, and why? How could AI alter the fundamental calculus of deterrence?
3. How could AI-assisted military platforms affect regional stability, and what is the connection between regional stability and strategic deterrence?
4. How will global competition in applying AI to military missions affect strategic stability? Should we be concerned about an "AI arms race"?
5. What are the risks of unintended consequences and strategic surprise driven by AI?

This paper frames large questions and provides first-order level arguments about them. It is intended to begin a conversation but not to delve systematically into any one specific aspect. It draws on ideas developed via workshops convened by Technology for Global Security and the Center for Global Security Research (CGSR) in June and September 2018, as well as ongoing

This project was a collaboration with the Center for Global Security Research at Lawrence Livermore National Laboratory (LLNL). The views expressed do not necessarily reflect those of LLNL, the Department of Energy, or the U.S. Government.

engagement with AI researchers based primarily in the California Bay Area.[5] These discussions have engaged a diverse mix of public and private sector experts in an exploration of the roles and consequences of AI in the military context. The paper also draws upon prior work at CGSR on disruptive and latent technologies in the 21st century security context.[6]

Defining "AI"

Much of the hype surrounding AI in the national security context stems from the fuzziness of our understanding of the technologies that combine to make what is now commonly referred to as "AI".[7] The term "AI" is used to describe a wide range of loosely-related techniques that are generally associated with using machines to glean insight from data in order to satisfy a set of goals. Similar to how the generic term "cyber" is used in reference to everything from networks to hardware, software, automation, industrial controls, hacking, bullying, warfare, and social media, AI is used as a generic term that washes over meaningful distinctions between different manifestations of AI. This breeds confusion, especially regarding claims about its revolutionary effects.

So far, the national security community lacks a detailed appreciation of the different technologies and the timelines by which they might mature into militarily-significant capabilities. For the vast majority of currently popular applications, AI consists of algorithms that form the basis of pattern recognition software that is able to interpret, organize, and make predictions. When combined with high performance computing power, data scientists are able to probe and find meaning in massive datasets. Neural networks supercharge the ability of the algorithms to identify and organize patterns in the data by "training" them to associate specific patterns with desired outcomes. Multiple layers of neural networks, known as deep learning neural networks, are what make current approaches to "machine learning," "supervised learning," and "reinforcement learning" possible.[8] However, the neural network approach represents only a fraction of the advancements in AI methods. For example, AI also includes language processing, knowledge representation, and inferential reasoning, which are all increasingly possible due to advancements in software, hardware, data collection, and data storage. Novel AI provides a leap in the ability to find needles in data haystacks–as long as you know what you are looking for.

Despite it being repeated often, it is useful to distinguish between "narrow" and "general" applications of AI. Narrow AI encompasses discrete problem-solving tools designed to perform specific, "narrow" tasks. General AI encompasses technologies designed to mimic and recreate functions of the human brain across domains. The gap between the two is significant. Most experts appear to agree that the accomplishments of narrow AI, though significant, are a long way from the requirements of replicating human-like reasoning. Although Alphabet's Deep Mind, OpenAI, and others have continued to make impressive breakthroughs in replicating human reasoning, they remain far from matching the performance of the human brain in its multiple contextualized and real-world dimensions. The quest for "superintelligence"[9] notwithstanding, recent progress in brain enhancement for now mostly replenishes impaired functions[10] and has a long way to go before it is possible to equip citizens, soldiers or robots with

super-human powers.[11] Although general AI stimulates intriguing science fiction about cyborgs, space wars, and robot armies, narrow AI is already here - and has been for some time.

Narrow AI is already in broad use in the private sector. In both business and science, AI has wide applications, primarily in data-rich research fields, including fundamental research (e.g., in physics, chemistry, and biology) and applied sciences (e.g., medicine, aeronautics, and environmental studies). Data science is facilitating rapid advancements in many aspects of scientific discovery, even changing long-held methodological standards and practices.[12] Figure 1 demonstrates scientific areas where AI-fueled deep learning is having its greatest effect.

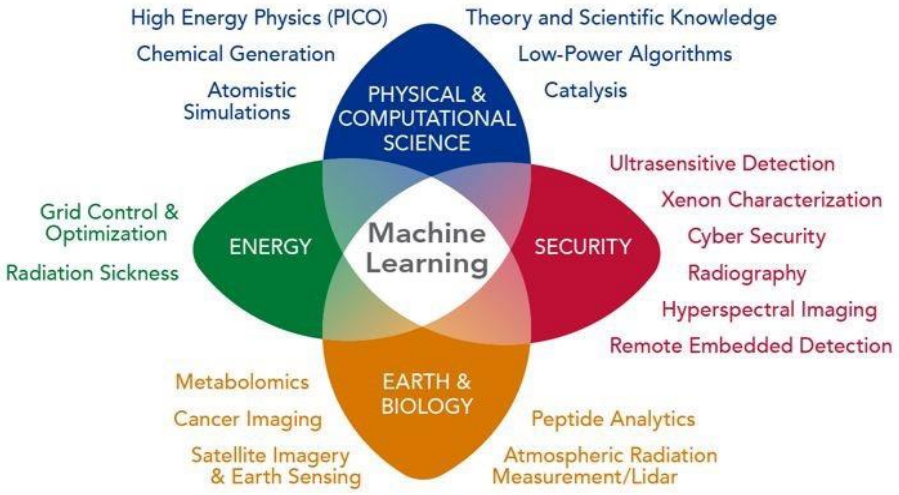


Figure 1. Disciplinary areas of deep learning for scientific discovery at the Pacific Northwest National Laboratory (PNNL). SOURCE: Nathan Hodas, PNNL, reproduced in National Academies Press, *Artificial Intelligence and Machine Learning to Accelerate Translational Research: Proceedings of a Workshop in Brief*, July 2018, available at http://nap.edu/25197.

The crossover of AI into business applications has super-charged predictive analytics for market research, consumer behavior, logistics, quality control, and many other data-rich areas. The proliferation of cameras and sensors creates even more opportunities for data analysis. When combined with robotics, AI is ushering in a new industrial age with far-reaching societal implications for labor, management, and economic equality.[13] For these scientific and business applications, AI is an enabling technology, a cross-cutting force multiplier when coupled with existing data-centric systems, such as the internet, health care, social media, industrial processes, transportation, and just about every aspect of the global economy, where recognizing patterns is the key to insight and profit. Growing interconnectivity, illustrated by the Internet of Things (IOT), is producing more data and providing more opportunity for AI algorithms to reveal hidden insights. For these types of applications, however, AI is increasingly more of a well-established, sustaining and enabling technology than a revolutionary new disruptive technology in its own right. Data analytics is not new, but it is getting better.

What Military Applications Are Possible?

With this context in mind, the focus of this paper is on AI applications in the military field. Like so many technologies, AI is loaded with latent military potential.[14] Algorithmic warfare is seen by many as the prime mover of a new revolution in military affairs.[15] AI is central to the so-called "third offset" strategy - originally coined by the U.S. Department of Defense in the second term of President Obama[16] - and thus has been a principal focus of multiple government initiatives to accelerate the development of advanced technologies. In June 2018, DOD established its Joint Artificial Intelligence Center[17] following the White House establishing its Select Committee on AI in May 2018.[18] DOD and IC spending on AI has also continued to increase.[19] The current U.S. Administration released the beginnings of a national strategy for AI in February, 2019.[20] For military applications with direct analogs in the civilian world, like logistics, planning, and transportation, AI-supported data analytics is already in use throughout the defense and intelligence communities.[21] These applications are separate and distinct from applications to warfighting. They tend to fall into one of two categories: those that have impact primarily at the operational level of war and those that have impact primarily at the strategic level of war. We define "strategic" as "extraordinarily consequential actions capable of causing a shift in the balance of power."[22] Examples of AI applications with impacts primarily at the operational level of war are:

- Omnipresent and omniscient autonomous vehicles
- Data driven modeling, simulation, and war-gaming
- Focused intelligence collection and analysis

Examples of AI applications with impacts primarily at both the operational and the strategic levels of war are:

- A system of systems approach enabling exquisite intelligence, surveillance, and reconnaissance (ISR)
- Delegation of decision-making to machines and the acceleration/improvement of decision support systems
- Automating integrated logistics systems for supply and maintenance
- Precision targeting of strategic assets
- Effective missile defense
- AI-guided offensive and defensive cyber operations

Omnipresent and Omniscient Autonomous Vehicles: Such vehicles are a high priority for military applications of AI, with much of the focus on navigation for unmanned land, sea and air systems.[23] AI is at the heart of the so-called "drone swarms" that have been the subject of much attention in recent years.[24] AI-informed navigation software supported by ubiquitous sensors not only enables unmanned vehicles to find their way through hostile terrain, but may eventually

make it possible for complex formations consisting of various types of unmanned vehicles operating in multiple domains, with complementary armaments to conduct sophisticated battle tactics, instantly adjusting to enemy maneuvers to rapidly exploit battlefield opportunities and report changing conditions.

Big Data-Driven Modeling, Simulation, and Wargaming: AI has steadily been increasing the power of simulations and gaming tools used to study nuclear and conventional weapons. From Samuel Glasstone's early calculations of nuclear effects, to the extensive library of RAND studies on nuclear issues, quantitative methods have been integral to the development of nuclear weapons systems. AI is enabling scientists to model nuclear effects to confirm the reliability of the nuclear stockpile without nuclear testing. Simulation and modeling is already a key part of the design process for nearly all major weapons systems, from jets and ships to spacecraft and precision guided munitions.[25]

Massive modeling and simulation will be necessary to design the all-encompassing multi-domain system of systems envisioned for battle management and complex missions such as designing, planning and managing systems for space situational awareness. On the production side, AI already informs quality control for novel production methods, such as additive manufacturing.[26]

AI is also enriching battlefield simulations and wargames involving multi-actor interactions. AI enables wargamers to add and modify game variables to explore how dynamic conditions (e.g., weapons, effects, allies, intervention) could affect outcomes and decision-making, and is used to analyze the results of such games. These are examples of evolutionary learning that are unlikely to cause strategic surprise or undermine stability unless the results negatively influence decision-making.

Focused Intelligence Collection and Analysis: With so many incoming streams of intelligence being collected from so many sources that all requiring analysis to be useful for policy maker (e.g., HUMINT, SIGINT, GEOINT, MASINT, ELINT, OSINT), the intelligence community faces the challenge of information overload.[27] This is a data-centric problem for which AI is well suited.[28] For example, a project at Lawrence Livermore National Lab (LLNL) uses neural networks to probe multi-modal data sets such as images, text, and video in search of key indicators of proliferation activity. AI also makes it possible to combine open source trade and financial data with multiple forms of intelligence to glean insights about illicit technology transfers, proliferation networks, and the efforts of proliferators to evade detection.[29] These insights enable analysts to more rapidly inform policymakers and support counterproliferation policy and actions.

AI will be an important tool for all-source analysts who are increasingly required to take into account information from many sources, locations, and disciplines to understand today's global security environment. To the extent that better information leads to informed decisions, applying AI to these collection and analysis problems would likely benefit strategic stability.

A System of Systems Enabling Exquisite ISR: For the military, object identification is a natural starting point for AI, as it requires culling images and data collected from satellites and drones to find information of military importance such as the location of missiles, troops, and other intelligence information. Accordingly, the National Geospatial Intelligence Agency (NGA) has led the charge in applying AI to military and intelligence needs.[30] But object identification is just the beginning. Intelligence, surveillance and reconnaissance (ISR) is the key to multi-domain situational awareness. This awareness is increasingly critical as the battlefield extends to all domains: sea, land, air, space, and cyber on a global scale.

Managing and making sense of the staggering amount of ISR data involved in modern warfare is a natural fit for AI, and the objective of DOD's Project Maven, also known as the Algorithmic Warfare Cross Functional Team, which has received significant increases in recent appropriations bills.[31] According to Lieutenant General Jack Shanahan, the Director of Defense Intelligence for Warfighter Support in the Office of the Under Secretary of Defense for Intelligence (and the incoming chief of the JAIC), Project Maven was conceived as "the spark that kindles the flame front for AI across the rest of the department."[32] While Maven's initial mission was to help locate ISIS fighters, its implications are vast.

Delegation of Decision-Making to Machines and the Acceleration/Improvement of Decision Support Systems: Advances in AI and access to large quantities of data represent a turning point in military decision-making. Multi-domain warfare involves colossal amounts of heterogeneous data streams that can only be exploited with the help of AI. Mirroring the proliferation of sensors in the civilian world, the multi-domain, hybrid warfare battlefield has become a military version of the Internet of Things, teeming with vital information for assessing tactical and strategic threats and opportunities. While the ability to manage this data colossus in real time portends tremendous advantages, failure to draw meaning from that information could spell disaster.

Being able to rapidly process the flood of information from varied platforms operating in multiple domains translates into two fundamental military advantages: speed and range. Moving faster than your adversary enhances offensive mobility and makes you harder to hit. Striking from further away similarly benefits the element of surprise and minimizes exposure to enemy fire. These were central tenets of the previous Revolution in Military Affairs that had its debut in the Gulf War. AI makes it possible to analyze dynamic battlefield conditions in real time and strike quickly and optimally while minimizing risks to one's own forces. As a recent Defense Science Board Study demonstrated, such integrated Battle Management, Command, Control, Communications, and Intelligence (BMC3I) capabilities are well suited to finding and targeting deployed missile batteries, and thus could be the key to countering critical elements of the Anti Access Area Denial (A2AD) strategies of Russia, Iran, and China.[33] These systems were designed to exploit vulnerabilities of U.S. land and sea assets in Europe and Asia. In addition to geo-locating targets, AI-enabled BMC3I could help guide and coordinate kinetic effects involving multiple platforms, possibly providing a counter to current A2AD. From this perspective, AI could be a strategic level game changer.

Automated Integration of Logistics Systems for Supply and Maintenance: AI technologies allow militaries to automate the coordination of logistical tasks across multiple platforms and locations, making it easier to run large, global operations with less resources and more precision. The ability to analyze massive datasets allows militaries to know when and where supplies are needed, predict when assets require maintenance, foresee supply chain disruptions, and integrate and coordinate several systems and logistical tasks. This in turn drives efficiency and allows for faster, more complex operations - mundane tasks perhaps, but as the adage goes, "Amateurs talk about tactics, but professionals study logistics." Private industry is already using AI to make faster and better decisions about supply chain and maintenance,[34] and the military is catching up with big investments in predictive maintenance.[35] These innovations will likely result in increased readiness, more seamless integration of assets across the global, and faster and more effective military strikes.

Precision Targeting of Strategic Assets: AI empowered ISR that makes it possible to locate, track and target a variety of enemy weapon systems raises the possibility of striking strategic assets, such as aircraft carriers, mobile missiles, or nuclear weapons. This capability, and perceptions of its existence, could disrupt long-held assumptions about deterrence stability, especially if it appeared possible to conduct a disarming counterforce strike against an adversary's retaliatory forces.[36] The combination of offensive weapons that can "find, fix and finish" a significant portion of an adversary's strategic assets,[37] with defensive systems that can shoot down remaining retaliatory capabilities, would challenge fundamental precepts of deterrence in the nuclear age.

Effective Missile Defense: Advancements in AI enhanced targeting and navigation also improve prospects for a wide range of tactical and strategic defense systems, especially ballistic missile defenses, by empowering target acquisition, tracking and discrimination.[38] The convergence of powerful new offensive and defensive capabilities has, however, rekindled fears of a surprise attack that could rattle strategic stability.

AI Guided Offensive and Defensive Cyber Operations: As an inherently digital domain, the cyber realm naturally lends itself to AI applications, as illustrated by the centrality of AI algorithms for social media companies. The availability of enormous amounts of data in electronic formats is well suited to AI strengths. AI-guided probing, mapping and hacking of computer networks can provide useful data for machine learning, including discovery of network vulnerabilities, identities, profiles, relationships, and other information that could be valuable for offensive and defense purposes.[39] On the offensive side, AI could help locate and target particular nodes or individual accounts for collection, disruption or disinformation. Cyber attacks on national command infrastructure and networks, for example, could be catastrophic.[40] On the defensive side of the equation, AI can help detect such intrusions and search for debilitating anomalies in civilian and military operating systems.[41] AI will equally empower offensive and defensive measures - as has been highlighted in DARPA's Cyber Grand Challenge.[42]

AI Threats to Strategic Stability

AI has multiple potential applications in the military domain, at both the operational and strategic level of war. But at the strategic level of war, some of the implications may not be altogether positive, as already foreshadowed above. Indeed, the disruptive effects of new technologies cannot be limited to the adversary. Some of those effects are potentially quite significant for the stability of strategic deterrence, by altering the fundamental calculus of deterrence. How might this be so?

In the classic Cold War movie *WarGames*, a young hacker breaks into a DOD supercomputer designed to use AI to plan and execute nuclear war plans. He engages the computer to play Global Thermonuclear War and accidentally triggers a simulated scenario of nuclear Armageddon, which is mistaken for the real thing. The computer ultimately learns that for nuclear deterrence, "the only way to win is not to play." If AI disrupts the central logic of nuclear deterrence, as understood by the nuclear powers, or fundamentally changes the underlying assumptions that support it, the strategic consequences could be far-reaching, and the prospects that computers will learn "not to play" uncertain.[43] There are several potentially destabilizing aspects of AI worth exploring:

- Increased Risk of War/Increased Risk of First Strike
- Deterrence and Strategic Parity
- Flawed Data
- Computer Vision
- Data Manipulation
- Ineffective Crisis Management
- Ineffective Joint Operations
- Unexpected Results
- Human-Machine Coordination
- Public Perception
- Accuracy in Decision-making
- Public Sector-Private Sector Tensions

Increased Risk of War: AI may be seen as eroding mutual strategic vulnerability and thereby as increasing the risk of war. The combination of exquisite ISR with an effective defensive shield could make it tempting to conduct a disarming, decapitating or blinding first strike at strategic targets, including nuclear command and control (NC3), early warning radars, or dual-capable

missiles and aircraft.[44] Such a revision of deterrence logic could be highly destabilizing. Shared vulnerability and assured retaliation are central concepts of mutually assured destruction (MAD) deterrence theory. Switching the theoretical incentive from MAD to improve the odds of successfully conducting a disarming first strike could change the risk calculus that has formed the basis of strategic stability for decades.[45] Preventing such a revision of nuclear deterrence logic was the essence of Vladimir Putin's claim in March 2018 that his new weapons are "invincible against all existing and prospective missile defense and counter-air defense systems."[46] By evading *perceived* U.S. global strike and missile defense capabilities, Putin's new AI-guided retaliatory forces are intended to preserve MAD. Additionally, if the weaker state believes a stronger adversary is actually positioned to strike first due to AI-fueled advantages (that will likely remain opaque and ambiguous due to their digital nature), they may want to lash out first in various ways (think Pakistan or DPRK).

Deterrence and Strategic Parity: No one country can gain all of the benefits of AI while denying them to potential adversaries. Competition to gain advantage will bring uncertainty about the future balance. Russia, China, and other nations' advancements in these same AI-enabled technologies has the potential to shift the strategic calculus as well, especially in regional contexts. For example, while Russian and Chinese A2AD systems designed to defeat U.S. regional forces may reduce U.S. allies' confidence in American security guarantees to protect them, the ability of the U.S. to defeat those A2AD systems with AI-accelerated ISR, BMC3I, defensive systems, and autonomous vehicles would demonstrate resolve and provide opportunities for joint U.S.-allied defense cooperation, thereby enhancing stability and deterrence. Reinforcing regional conventional deterrence is also an essential part of strategic stability.[47] However, even the *perception* of an imbalance that favors striking first can lead to misperception, miscalculation, and arms racing. Whatever advantages can be attained with AI are likely to evoke countermeasures that mitigate temporary unilateral advantages, with all the incumbent opacity and AI safety risks referred to above.

Flawed Data: AI systems are vulnerable to flawed data inputs. In her book *Weapons of Math Destruction*, data scientist, Cathy O'Neil, demonstrates how AI algorithms distort reality and lead to incorrect, misleading, and unjust decisions.[48] Perhaps the biggest obstacle to increasing reliance on AI is the age-old problem of data reliability. AI can magnify the "garbage in, garbage out" problem.[49] Data comes from many places and is not always carefully collected or curated, particularly as industry-driven innovation competes in a race-to-the-bottom for market share. Compounding the problems with the data itself leading to skewed results, AI often reflects human bias,[50] or creates new biases based on flawed "learning" from the data provided.[51]
Computer Vision: The AI-informed object and pattern recognition software behind Project Maven and many other applications - is relatively easily fooled by misleading data.[52] Differentiating between similar objects is difficult,[53] and more challenging with denial and deception campaigns, such as the use of camouflage and decoys. Even when data seems accurate, AI sometimes "hallucinates" things that do not exist.[54] Transferring these inherent problems of data reliability and interpretation onto the battlefield raises critical questions about

the safety and reliability that comes with the desirable qualities of speed and lethality. Accidentally hitting the wrong targets, for example, could have strategic consequences.

Data Manipulation: Countering many AI applications can be simple and straightforward. Adversarial manipulation of data (image, text, etc.) provides many opportunities for mischief and mistakes.[55] The fact that AI is easily deceived invites efforts to counter the sought-after military benefits.[56] By corrupting data in calculated ways, it may be possible to cause catastrophic equipment failures, miscommunication, confusion, logistical nightmares, and devastating mistakes in AI-reliant systems. Additionally, these challenges can be compounded by the "black box" problem of not understanding how and why AI makes decisions also means that it would be hard to recognize if data had been compromised to produce inaccurate outcomes, such as hitting the wrong targets or misdirecting U.S. and allied forces.

Ineffective Crisis Management: Speedy decision-making and operational execution may not serve well the goals of effective crisis management. On October 19, 1962, only three days into the Cuban Missile Crisis, General Curtis LeMay counseled President Kennedy, "I just don't see any other solution except direct military action right now." [57] Ten days later, the crisis was resolved diplomatically. If one of the advantages of AI is the speed it adds to decision-making, that same speed could be a disadvantage if it accelerates the escalation of conflict from crisis, to war, and even potential nuclear confrontation.[58] The battlefield advantages of AI-driven ISR and autonomous systems could shrink the time available for diplomacy to avoid or manage crises. As currently conceived, AI-driven battlefield systems would not include real time reporting and analysis of national and international diplomatic efforts to avoid, control, contain, or end a conflict - violating Clausewitz's principle of war as "the continuation of politics by other means." In many cases, logic might dictate striking first, as General LeMay advised. Accelerated decision-making might have pushed the Cuban Missile Crisis toward dangerously different outcomes. In practice, slowing things down can be the key to victory, especially when the stakes involve nuclear weapons. The fact that this accelerated information in the hands of decision makers is likely to be derived from insecure and unsafe AI-based systems even further exacerbates the command and control conundrum: how much risk and what systemic error rates are acceptable when contemplating nuclear war, and what are the signaling mechanisms in this theoretical hyper-speed future?

Ineffective Joint Operations: AI may be unhelpful in trying to solve the main problem confronting U.S. military forces: effective joint operations. Such operations are always challenging, even for the most advanced military in human history. AI-supported weapons, platforms, and operating systems operate according to custom-built software and hardware that is specifically designed for each separate problem, system, and purpose. There is currently no overarching mechanism to integrate scores of AI-powered systems operating on multiple platforms.[59] To get the desired effects of multi-domain ISR, it is necessary to integrate across scores of sensors, radars, weapons, and communications systems operating in multiple geophysical domains. If this were not challenging enough, those systems would be built and operated by different agencies, commands, and contractors, with different authorities,

procedures, and operational cultures. Adding allies with their own AI systems to this landscape brings further complexity and risk.

Designing a multiplex of AI-informed platforms that have the ability to communicate with one another in real time requires a new generation of data fusion, integrative software, and command architectures. Moreover, pulling all these pieces together to develop a holistic threat assessment that provides policymakers with strategic warning will not happen naturally. Instead, this task will require Herculean efforts to collect and analyze information "owned" by diverse stakeholders. Inter-service rivalry when it comes to AI has already begun to take root, and has severely complicated efforts to streamline the acquisitions, technology, and logistics process within the U.S. Defense Department. It remains unclear how these internal cultural challenges may play out in a more authoritarian environment such as Russia or China. Developing a fully integrated system capable of providing strategic warning in the United States, however, will clearly take many years.

Unexpected Results: The close operation and integration of multiple AI systems, as required on the battlefield, can be expected to have unexpected results. The flip side of stovepiped systems not talking to each other is the issue of unexpected convergences. It is uncertain how separate AI-infused platforms might interact with one another, as various AI-guided systems operate in shared battlespace. Unknown outcomes resulting from friendly interactions are likely to be compounded by interactions with foreign AI systems. With so much uncertainty about the internal "black box" mechanisms that produce AI outcomes, AI to AI interactions are likely to produce unanticipated and unexplainable results, like choosing the wrong targets.[60] Lastly, we cannot anticipate how AI will converge with other technologies, such as quantum computing, electromagnetic pulses, IoT, 5G, or blockchain/distributed ledgers. Potential convergences could produce strategic surprises that confuse and confound friends and foes alike, making the fog of war even more impenetrable.

Human-Machine Coordination: Whether or not there is a human in the loop, the loop is getting crowded. The interface between humans and machine–where the proverbial "man in the loop" is supposed to exert human control–also raises critical questions about decision-making authority and organizational hierarchies.[61] Within the military, questions of rank, service branch, and responsibility for lethal actions can be contentious in the best of times, as illustrated by the debates over authority for U.S. drone strikes.[62] With scores of AI-informed battlefield systems operating at breakneck speed, each connected to its own chain of command, coordination among the humans who are in the loop of fast-moving battlefield operations spanning multiple adversaries, domains, agencies, clearance levels, contractors, allies, and organizational cultures will be challenging, especially if the goal is to maintain offensive advantage via speedy decision-making. Budgets, reorganizations, access, personalities, and leadership changes may have as much influence over AI capabilities as the technology itself. There will be lots of people in the loop in lots of places, each influencing how AI contributes to separate and shared objectives. Achieving strategic effects will require extraordinary cooperation and communication—again, a complication that a more authoritarian adversary may or may not also be grappling with.

Public Perception: It is hard to predict how the public will respond to AI in the military**.** AI algorithms are a central component of cyber influence operations aimed at shaping public perceptions. By now it should be understood that the use and misuse of electronic media to manipulate public perceptions, including the use of fake news, cyber bots, and deep fakes, can affect strategic stability.[63] How the public views particular international conflicts can shape leadership decision-making, and can build or undermine support for issues of war and peace, especially in democratic states. AI powered tools such as cyber bots and deep fake technology could enrage or pacify public opinion, or mislead decisionmakers. Now that cyber conflict has become an ingrained feature of the international landscape, we should expect AI-fueled manipulation of public perceptions to affect crisis management, escalation, deterrence stability, and possibly nuclear decision-making.

Accuracy in Decision-making: Decisions of war and peace cannot be left to predictive analytics. There are fundamental differences in the ways that data is used for scientific, economic, and logistic purposes, and for predicting human behavior. AI still cannot reliably predict the outcomes of sports contests, elections, or international conflict, at least within acceptable margins of error for making consequential decisions. Despite longstanding interest in predictive analytics that can tell decisionmakers what to expect before it happens, faith in the ability to predict incidence or outcomes of war and conflict based on big data machine learning is fraught with misplaced optimism.[64] All of the potential dangers stemming from unreliable (outdated, biased, compromised) data, machine learning bias, and interpretation errors are magnified when human emotions, non-rational behavior, and inherent unpredictability cloud the data *and* the decisionmaking. The result is wider margins of error which may be acceptable for academic purposes, but do not satisfy practical and ethical demands of national security decisionmaking.

Much like self-driving cars, where AI can correctly assess most–but not all–situations, 90% correct predictions could mislead decision makers and put soldiers and citizens lives at stake. Close is not good enough when it comes to war, especially where nuclear risks are involved. Nation states, however, are moving forward with integrating these AI-techniques into their militaries and command structures. While the United States, China, and Russia move to integrate these technologies into their militaries, there remains a paucity of research and understanding regarding the cumulative risk inherent to large decision support systems that rely on these technologies not just to potentially provide predictions and recommendations–but also will be responsible for an increasingly large percentage of the information provided to decision-makers through the command and control enterprise. While flawed predictions may present an unacceptable scenario with horrific potential outcomes, an entire ecosystem built on unsafe and unpredictable AI-fueled systems quite simply remains an un-investigated and under-appreciated element of the potential risks inherent for strategic stability and nuclear deterrence.

Public Sector-Private Sector Tensions: public-private partnerships shape the future of AI, but war remains the preserve of the state. As a dual-use technology, AI is freely available to everyone. It is being developed and applied beyond the reach of governmental controls. Like many other

dual-use technologies, governments rely on the private sector for the underlying research and development, software, hardware and expertise required for AI to be used for military purposes. DOD and the intelligence community have deep ties to Silicon Valley and have doubled down on efforts to expedite the acquisitions process, especially for cyber and AI.[65] What this means in practice is that many countries will use the same experts, companies, and global supply chains to support their military AI aspirations, creating potential competitive conflicts of interest and security vulnerabilities related to sharing intellectual property. This dynamic is already evident in the global marketplace, where Google and other companies have found it advantageous to accommodate Chinese demands on censorship and surveillance,[66] while simultaneously expressing political opposition to supporting U.S. military AI projects such as Project Maven. Global technology companies will have to weigh the costs and benefits of serving some customers while keeping others at arm's length. The U.S. government, however, has little choice but to remain heavily dependent on the private sector to develop and implement AI strategies, while also competing with adversaries who are not constrained by the separation between public and private. China, for example, continues to move forward with haste.[67]

AI's Potential Effects on Deterrence

With these potential destabilizing effects in mind, how could AI alter the fundamental calculus of deterrence, ?

At the top of the list of AI applications that could have true strategic significance is the likely increased risk of surprise attack. The combination of effective defenses with exquisite ISR that makes it possible to locate mobile targets and strike them with speed and precision raises long-held fears of a "bolt from the blue" first strike. While the fundamental logic of deterrence is unchanged, perceptions that an adversary has sufficient intent and capability to conduct such a preemptive attack on vital assets motivates a variety of counter measures.

Evaluating the incentive to strike first evokes memories of Pearl Harbor, in which the U.S. underestimated Japan's risk calculus while fully recognizing Tokyo's military capacity to launch a cross-Pacific raid. AI contributions to military and intelligence capabilities do not override political considerations - with an important caveat added for the possibility of AI-fueled manipulation of public attitudes that could distort political judgement. Avoiding and deterring conflict remains a paramount responsibility for national leaders. Slightly improved odds of eliminating all but a few of an adversary's strategic weapons, and shooting down any surviving retaliation with missile defenses, still involves catastrophic risks, and does not even begin to answer questions about the aftermath of such a conflict.

Nevertheless, possessing the theoretical capability to conduct a disarming first strike inevitably triggers a classic security dilemma, which is guaranteed to provoke counter measures from those threatened by enhanced striking power. Further advances in defenses against counterforce strikes would be a predictable response, as well as hardening and camouflage to evade and confuse exquisite ISR. Asymmetric capabilities to undermine said advantages also will arise, with cyber

an immediate domain for consideration. To the extent that AI influences perceptions of intent and capability, and alters the calculus of risk and reward, it will inspire new thinking about possible offensive and defensive maneuvers in the Evolution of Nuclear Strategy.[68]

AI and Regional Stability

How could AI-assisted weapon systems affect regional stability, including U.S. allies?

Widespread deployment of AI-supported ISR platforms is likely to affect regional stability in the five to ten year time frame. While the United States remains the leader in translating AI to currently deployed platforms, China and Russia are not far behind.[69] Many U.S. allies are advancing their own AI capabilities, albeit with significant constraints. Initially, however, the speed and lethality gained from AI-informed situational awareness and battle management systems is likely to provide the United States and its allies with options for countering Russian and Chinese A2AD. The coming architecture of ISR, BMC3I, and defensive systems appears well positioned to give net advantages for U.S. and allied regional security alliances. In addition to tactical military benefits, co-development of multi domain ISR provides opportunities for collaboration that directly address threats to allied security, especially with respect to extended deterrence relationships with key allies in Asia and Europe. Strengthening regional conventional deterrence and regional extended nuclear deterrence reduces incentives for risk taking and supports broader interests in strategic deterrence. AI applications that support these objectives will have beneficial effects for strategic stability.

Strategic AI Competition, Arms Racing, and Strategic Stability

How will competition in applying AI to military missions affect strategic stability?

Global competition in military AI continues to intensify. A general consensus is forming indicating that an AI arms race is underway. Whatever advantages are possible in the near term, however, may be short lived, as U.S. allies, major adversaries, and a multitude of rising powers incorporate AI into their political and military strategies. In light of the rising tide that is advancing AI prospects around the world, temporary advantages are unlikely to yield lasting military predominance. For example, China and Russia will eventually possess their own versions of multi domain ISR coupled with precision strike and layered defenses. How will these capabilities influence Beijing's thinking about the U.S. role in the South China Sea, or Russian assessments of NATO's defense of the Baltics? These are not primarily technical issues. AI is enhancing the performance of many tactical and strategic systems, but not giving definitive unilateral advantage to anyone. The conduct of warfare is changing, and AI is fueling many of those changes, but it remains to be seen if the calculus of deterrence will remain steady. While competition that retains a balance of power can be stabilizing, the accompanying uncertainty has historically led to miscalculation and great power conflict.

Unintended Consequences and Strategic Surprise

What are the risks of unintended consequences and strategic surprise from AI?

Predicting the future impact of technology is a risky business. We know with certainty that AI is being incorporated into a wide array of military missions with the intent of improving our knowledge of the operational environment, an adversary's capabilities, and the speed and precision of offensive and defensive weapons. We can usefully speculate about how these developments are poised to change the face of modern warfare and how those changes might affect regional and strategic deterrence stability, based on our understanding of established political and military realities. More elusive, however, is a clear picture of how AI might converge with other technologies to produce unexpected outcomes or "unknown unknowns." Nevertheless, several possibilities that could have major strategic consequences and alter the underlying realities on which regional and strategic stability are founded:

- Distorted data could lead AI systems to take unintended actions or provide inaccurate analytical predictions/recommendations, such as incorrectly identifying and striking the wrong targets. As discussed above, data can be polluted intentionally via counter AI methods, or occur technically for many reasons that are yet to have been solved even by leading global AI researchers. Such actions could hasten escalation and/or interfere with conflict management efforts.

- Compounding the problems of distorted data, AI makes mistakes with a frequency that could be untenable for decisions affecting strategic stability. Misinterpretations of data that lead to unintended actions could spark highly undesirable reactions, including escalation and retaliation.

- The convergence of AI and cyber presents several possibilities for unintended consequences and strategic surprise. AI-informed cyber attacks on NC3 could present the target of such an attack with a "use it or lose it" situation, prompting early resort to nuclear weapons - potentially even targeting the wrong actors.

- AI supported cyber/information warfare, including use of dis/mis-information, machine-derived false media, and other methods could distort public and leadership perceptions of international events, inflaming passions and prompting escalation, or even accidentally triggering nuclear early warning systems.

- Accelerated battle rhythm made possible by multi-domain ISR could preclude diplomatic efforts to avoid or de-escalate conflict. Even if AI works perfectly to increase the speed and lethality of warfare, moving at the speed of AI might not be optimal for all cases.

- Unpredictable AI interactions with foreign and friendly platforms could produce unwanted AI calculations that misrepresent human intentions. The "black box" underlying AI decisions is not well understood and could produce destabilizing results, such as striking the wrong targets, or striking targets that posed no actual threat and instigating escalatory military actions.

- Unexpected convergences with other technologies, such as quantum computing and electromagnetic pulse, could confuse/distort offensive or defensive instructions and lead to undesirable results, such as striking the wrong targets.

- If it were eventually possible through a variety of AI-supported information gathering methods, emerging technologies, and analytic tools to track strategic assets such as submarines, the sanctity of assured retaliation could come into question. Such a strategic surprise could prompt a variety of destabilizing actions, including possible movement toward launch on warning postures.

Questions to Consider Moving Forward

This paper covers a wide range of topics without delving deeply into any particular question. Further research is required to fully comprehend the potential risks and rewards of AI in military operations. They include:

1. What are the implications of integrating AI into military decision-making and decision support systems?
2. How risky are the vulnerabilities inherent to AI-related technologies in the military context, and are we able to tolerate or eliminate those risks?
3. When designing and utilizing an AI system, when do you trust the human over trusting the machine?
4. How will AI empower unstable or non-state actors with few resources?
5. How do we successfully and safely integrate, verify, and validate AI tools in actual military operations?
6. Can we ensure the accuracy/security of data sets?

What are the implications of integrating AI into military decision-making? The national security community is beginning to understand the capabilities associated with novel AI-techniques, but there is very limited understanding of the compounding risks posed by the inherent safety and security challenges that these novel techniques face. Research remains necessary on examining the implications of an entire decision support system that is providing advice to a military commander that cannot actually be understood. Further research also remains necessary about what is considered acceptable error and risk thresholds, and whether/how these systems compound risk, or how they may be distinct.

How risky are the vulnerabilities inherent to AI-related technologies in the military context, and are we able to tolerate or eliminate those risks? Several instances already exist in the private sector of AI-related technologies exhibiting bias or failure. While the harm created by faulty chatbots and self-driving cars can be contained, how and to what extent can AI create problems in the military context, especially when critical military decisions are made based on AI systems or even made by AI systems? Do militaries have adequate safeguards for AI errors, and are such safeguards even possible given the speed and pressures of warfare? How will the military handle primary safety and security challenges of AI, namely issues like adversarial examples, data poisoning, reward hacking, and data protection? Are military leaders providing adequate guidance to personnel on how to use and safeguard AI systems?

When designing and utilizing an AI system, when do you trust the human over trusting the machine? Many people are concerned with the "singularity" where AI surpasses human cognition. However, this is not how AI currently works. Cognition isn't a straight line but rather a complex, multidimensional process. Are there cognitive functions or combinations of functions where humans will always be superior? In areas where AI might surpass human cognition, like deductive reasoning, how and when do we decide "trust" its decisions within a military system? When should a human be "in-the-loop," "on-the-loop," and "out-of-the-loop?" Will humans be pushed out of decision-making altogether? Additionally, super-human performance associated with AI likely means counter-intuitive decisions. How would do react to those decisions? How well can humans trust machines that cannot explain their recommendations - especially in a military context? Scenario-based exploration of these questions will illuminate new answers.

Can AI empower unstable or non-state actors with few resources? AI allows militaries to perform labor-intensive work with much less manpower and at significantly lower costs. It also gives militaries the ability to perform more tasks with less resources. But are these same benefits already lowering barriers to entry for unstable regimes or non-state actors to build their own AI capabilities, and will they do so in the future?

How Do We Successfully and Safely Integrate AI into Military Operations? Integrating AI-related technologies into military operations requires training, hiring technology-savvy people on the ground, and ensuring coordination and interoperability between multiple systems, products, and technology across the world. What can we do to ensure that AI-related technologies are safely integrated into military operations given these implementation challenges?

Can we ensure the accuracy/security of data sets? Furthermore, algorithms and deep-learning tools are dependent on clean, accurate data sets to train or improve their capabilities. But there have already been instances where algorithms, bots, and data sets were hacked or poisoned, and there would be dire consequences in the military setting where such attacks would impact troop movement, weapons deployment. What kinds of safeguards and precautions need to be in place to prevent such incidents, and, again, are such safeguards even possible?

Conclusion

Evolutionary changes in the logic of regional and strategic deterrence are not new, nor are they necessarily harmful to U.S. national security. Efforts to integrate AI-based technologies into U.S. defense and intelligence strategies illustrates the continued innovation and competitive advantages sought in support of U.S. national security policy. Recent changes in U.S. deterrence posture may not end up undermined by these developments in AI, at least in the near term.[70] However, the rapid expansion of AI military applications across the globe merits a high level of focused attention to minimize its negative impacts on strategic stability, and to prevent strategic surprise. As states and non-state actors being to deploy a broad range of AI-techniques in weapons systems, command and control, and decision support systems, the fundamental assumptions underlying survivability and credibility may well be called into question over time. As with previous eras of significant evolutionary change in the conduct of warfare, prudent planning and novel strategic frameworks combined with intense diplomatic engagement will be required to prevent destabilizing effects from these novel techniques. Finally, a deep integration of verification and validation of the safety and security of AI systems must be a pillar of planning, programming, budgeting and execution of defense acquisition processes going forward. Without this in place, the vulnerabilities created by these new technologies in and of themselves could lend to significant destabilizing trends. Adding additional machine fallibility to human imperfection seems a perhaps unwise endeavor, and will require greater systematic research and understanding of the implications of these novel techniques in relatively short order.

BIBLIOGRAPHY

[1] Anthony Cuthbertson, "What's Bigger Than Fire and Electricity? Artificial Intelligence, Says Google Boss," *Newsweek*, January 22, 2018, https://www.newsweek.com/artificial-intelligence-more-profound-electricity-or-fire-says-google-boss-786531; "Elon Musk: Mark my words – AI is far more dangerous than nukes" remarks at SXSW, CNBC, March 13, 2018, https://www.cnbc.com/2018/03/13/elon-musk-at-sxsw-a-i-is-more-dangerous-than-nuclear-weapons.html; Peter Holley, "Stephen Hawking just got an artificial intelligence upgrade, but still thinks AI could bring an end to mankind," *Washington Post*, December 2, 2014.

[2] "Whoever leads in AI will rule the world: Putin to Russian children on Knowledge Day," September 1, 2017, RT News, https://www.rt.com/news/401731-ai-rule-world-putin/. Accessed July 31, 2018.

[3] Paul Mozur, "Beijing Wants A.I. to be Made in China by 2030," *The New York Times,* July 20, 2017, https://www.nytimes.com/2017/07/20/business/china-artificial-intelligence.html Accessed September 24, 2018; Carlos Perez, "The West in Unaware of The Deep Learning Sputnik Moment," September 10, 2017, https://medium.com/intuitionmachine/the-deep-learning-sputnik-moment-3e5e7c41c5dd.

[4] Patrick Tucker, Samuel Bendett, Elsa Kania, and Josh Kirshner, *The Race for AI: The return of great power competition is spurring the quest to develop artificial intelligence for military purposes*, Defense One ebook, March 2018, https://www.defenseone.com/assets/race-ai/portal/.

[5] Summaries are available on the T4GS and CGSR websites, along with an annotated bibliography aligned with the workshop agenda. See https://cgsr.llnl.gov/ and www.tech4gs.org.

[6] See Zachary Davis, et. al., https://cgsr.llnl.gov/content/assets/docs/Strategic_Latency.pdf.

[7] While some in the technical community have argued that "AI" as a classification of these technologies is misleading and confusing enough that it should be avoided entirely - instead arguing specific techniques should be explicitly named in each instance, for example deep reinforcement learning, etc. For the purposes of this first-order level analytical paper, however, we will use AI in the broad sense that it is commonly understood and deployed.

[8] Jürgen Schmidhuber, "Deep Learning in Neural Networks: An Overview," Neural Networks, Volume 61, January 2015, pp. 85–117, https://doi.org/10.1016/j.neunet.2014.09.003.

[9] As defined by Nick Bostrom, superintelligence is "an intellect that is much smarter than the best human brains in practically every field, including scientific creativity, general wisdom and social skills". Widely seen as the seminal work on "superintelligence" is his https://nickbostrom.com/superintelligence.html.

[10] Sara Reardon, AI-controlled brain implants for mood disorders tested in people, *Nature*, November 22, 2017; Antonio Regalado, "Reversing Paralysis: Scientists are making remarkable progress at using brain implants to restore the freedom of movement that spinal cord injuries take away," *MIT Technology Review*, https://www.technologyreview.com/s/603492/10-breakthrough-technologies-2017-reversing-paralysis/.

[11] Sara Reardon, The Pentagon's gamble on brain implants, bionic limbs and combat exoskeletons," *Nature*, June 10, 2015; Annie Jacobsen, Engineering Humans for War, *The Atlantic*, September 23, 2015, https://www.theatlantic.com/international/archive/2015/09/military-technology-pentagon-robots/406786/; Michael Joseph Gross, "The Pentagon's Push to Program Soldiers' Brains," *The Atlantic*, November 2018, https://www.theatlantic.com/magazine/archive/2018/11/the-pentagon-wants-to-weaponize-the-brain-what-could-go-wrong/570841/.

[12] David Weinberger, "Our Machines Now Have Knowledge We'll Never Understand," *Wired*, April 18, 2017, https://backchannel.com/our-machines-now-have-knowledge-well-never-understand-857a479dcc0e.

[13] Darrell West, *The Future of Work: Robots, AI and Automation*, (Washington: Brookings Institution Press, 2018); Molly Kinder, "Learning to Work With Robots: AI will change everything. Workers must adapt – or else," *Foreign Policy*, July 11, 2018, https://foreignpolicy.com/2018/07/11/learning-to-work-with-robots-automation-ai-labor/.

[14] Zachary Davis and Michael Nacht, eds., *Strategic Latency: Red, White and Blue, Managing the National and International Security Consequences of Disruptive Technologies*, (Livermore, CA: Lawrence Livermore National Laboratory, 2018) Available at CGSR.LLNL.gov.

[15] F.G. Hoffman, "Will War's Nature Change in the Seventh Military Revolution?" Exploring War's Character and Nature, Parameters, 47(4) Winter 2017-18.

[16] Deputy Secretary: Third Offset Strategy Bolsters America's Military Deterrence, Department of Defense, October 31, 2016, https://dod.defense.gov/News/Article/Article/991434/deputy-secretary-third-offset-strategy-bolsters-americas-military-deterrence/.

[17] Memorandum from the Deputy Secretary of Defense, "Establishment of the Joint Artificial Intelligence Center," June 27, 2018.

[18] *Summary of the 2018 White House Summit on Artificial Intelligence for American Industry*, The White House, May 10, 2018.

[19] DARPA Announces $2 Billion Campaign to Develop Next Wave of AI Technologies, DARPA, September 7, 2018, https://www.darpa.mil/news-events/2018-09-07.

[20] Michael Kratsios, *Why the U.S. Needs a Strategy for AI*, The White House Office of Science and Technology Policy, February 12, 2019, https://www.whitehouse.gov/articles/u-s-needs-strategy-ai/.

[21] Daniel Hoadley, Nathan Lucas, *Artificial Intelligence and National Security*, Congressional Research Service, April 26, 2018; Marcus Weisgerber, The Pentagon's New Artificial Intelligence Is Already Hunting Terrorists," Defense One, December 21, 2017, https://www.defenseone.com/technology/2017/12/pentagons-new-artificial-intelligence-already-hunting-terrorists/144742/; Matt Leonard, "Army leverages machine learning to predict component failure, Defense Systems, July 2, 2018, https://defensesystems.com/articles/2018/07/03/army-vehicle-predictive-maintenance.aspx.

[22] *Strategic Latency and Warning: Private Sector Perspectives on Current Intelligence Challenges in Science and Technology*, Report of the Expert Advisory Panel Workshop, Lawrence Livermore National Laboratory, January 8, 2016. *Strategic warning* describes the goal of alerting decision makers of impending threats of a strategic nature. *Strategic surprise* describes the failure to provide adequate warning of such threats.

[23] National Academies of Science, *Autonomy in Land and Sea and In the Air and Space, Proceedings of a Forum*, 2018, http://nap.edu/25168.

[24] National Academy of Sciences, *Counter-Unmanned Aircraft System (CUAS) Capability for Battalion and Below Operations, Abbreviated Version of a Restricted Report*, 2018, http://www.nap.edu/read/24747/chapter/1.

[25] Lisa Owens Davis, "Moving at the Speed of S&T: Calibrating the Role of National Laboratories to Support National Security," in Davis and Nacht, *Strategic Latency: Red, White and Blue*, ibid.

[26] "Machine learning to prevent defects in metal 3D printed parts in real time," Lawrence Livermore National Lab, *Newsline*, September 13, 2018, https://webcenter.llnl.gov/myllnl/faces/oracle/webcenter/portalapp/pages/top-story-wrapper.jspx?articleId=52535&_afrLoop=77869951013468&_afrWindowMode=0&_afrWindowId=blank#%40%3F_afrWindowId%3Dblank%26_afrLoop%3D77869951013468%26articleId%3D52535%26_afrWindowMode%3D0%26_adf.ctrl-state%3Dt66qlfya5_65.

[27] Marc Pomerleau, "Can the intel and defense community conquer data overload? C4ISRNET, September 5, 2018, https://www.c4isrnet.com/intel-geoint/2018/09/05/can-the-intel-and-defense-community-conquer-data-overload/?utm_source=Sailthru&utm_medium=email&utm_campaign=daily%20brief%209/5/18&utm_term=Editorial%20-%20Daily%20Brief.

[28] Marc Pomerleau, "Here's how intelligence agencies will take advantage of machine learning and AI," C4ISRNET, May 1, 2018, https://www.c4isrnet.com/intel-geoint/2018/05/01/heres-how-intelligence-will-take-advantage-of-machine-learning-and-ai/.

[29] "Deep Learning to Advance Nuclear Nonproliferation," *LLNL Newsline*, August 21, 2018, https://webcenter.llnl.gov/myllnl/faces/oracle/webcenter/portalapp/pages/top-story-wrapper.jspx?articleId=52206&_afrLoop=17092740707130&_afrWindowMode=0&_afrWindowId=null#%40%3F_afrWindowId%3Dnull%26_afrLoop%3D17092740707130%26articleId%3D52206%26_afrWindowMode%3D0%26_adf.ctrl-state%3D1af49b8608_69.

[30] Ben Conklin, "How artificial intelligence is transforming GEOINT," *GCN*, April 18, 2018, https://gcn.com/articles/2018/04/18/ai-transform-geoint.aspx; Sandra Erwin, "NGA official: Artificial intelligence is changing everything, We need a different mentality," *Spacenews*, May 13, 2018, https://spacenews.com/nga-official-artificial-intelligence-is-changing-everything-we-need-a-different-mentality/.

[31] Kelsey Atherton, "Targeting the future of the D0D's controversial Project Maven initiative," C4ISRNET, July 27, 2018, https://www.c4isrnet.com/it-networks/2018/07/27/targeting-the-future-of-the-dods-controversial-project-maven-initiative/.

[32] Jack Corrigan, "Project Maven uses machine learning to go through drone video feeds, but that just the beginning, Air Force Lt. Gen Shanahan said," *Nextgov*, November 2, 2017, https://www.nextgov.com/cio-briefing/2017/11/three-star-general-wants-artificial-intelligence-every-new-weapon-system/142225/.

[33] Defense Science Board, *Study on Countering Anti-access Systems with Longer Range and Standoff Capabilities: Assault Breaker II*, 2017 Summer Study on Long Range Effects, Office of the Under Secretary of Defense for Acquisition, Technology, and Logistics, June 2018.

[34] Forbes Insights, *Logistics, Supply Chain and Transportation 2023: Change at Breakneck Speed*, 2018, available at https://www.forbes.com/forbes-insights/our-work/transportation-2023/.

[35] Sonja Jordan, *Army Investing in Predictive Maintenance for Bradleys,* National Defense, Sept. 26, 2018, available at http://www.nationaldefensemagazine.org/articles/2018/9/26/army-investing-in-predictive-maintenance-for-bradleys.

[36] Edward Geist and Andrew Lohn, *How Might Artificial Intelligence Affect the Risk of Nuclear War?* RAND, 2018; Paul Bracken, "The Intersection of Cyber and Nuclear War," The Strategy Bridge, January 17, 2017, https://thestrategybridge.org/the-bridge/2017/1/17/the-intersection-of-cyber-and-nuclear-war.

[37] Jeremy Hsu, "AI Can Help Hunt Down Missile Sites in China," *Wired,* November 21, 2017, https://www.wired.com/story/ai-can-help-hunt-down-missile-sites-in-china/.

[38] Jen Judson, "Hyten: To address Russian and Chinese missile threats, it's all about the sensors, " Defense News, August 7, 2018, https://www.defensenews.com/digital-show-dailies/smd/2018/08/07/hyten-to-address-russian-and-chinese-missile-threats-its-all-about-the-sensors/.

[39] Jack Corrigan, "DARPA Wants to Find Botnets Before They Attack," *Defense One*, September 12, 2018, https://www.defenseone.com/technology/2018/09/darpa-wants-find-botnets-they-attack/151199/?oref=defenseone_today_nl.

[40] Nuclear Weapons in the New CyberAge: A Report of the Cyber-Nuclear Weapons Study Group, Nuclear Threat Initiative, September 2018, https://www.nti.org/media/documents/Cyber_report_finalsmall.pdf.

[41] Michael Sulmeyer and Kathryn Dura, Beyond Killer Robots: How Artificial Intelligence Can Improve Resilience in Cyber Space, Sept 6, 2018, *War on the Rocks*, https://warontherocks.com/2018/09/beyond-killer-robots-how-artificial-intelligence-can-improve-resilience-in-cyber-space/.

[42] Dustin Fraze, DARPA Cyber Grand Challenge Program Information, https://www.darpa.mil/program/cyber-grand-challenge.

[43] Recent related work for consideration in this vein includes Richard Danzig, "Technology Roulette: Managing Loss of Control as Many Militaries Pursue Technological Superiority", May 2018, https://www.cnas.org/publications/reports/technology-roulette; and Michael Horowitz, "Artificial Intelligence, International Competition, and the Balance of Power", May 2018, https://tnsr.org/2018/05/artificial-intelligence-international-competition-and-the-balance-of-power/).

[44] James Acton, "Escalation through Entanglement: How the Vulnerability of Command and Control Systems Raises the Risks of an Inadvertent Nuclear War," *International Security*, Volume 43, Summer 2018.

[45] Kier Lieber and Daryl Press, "The New Era of Counterforce: Technological Change and the Future of Nuclear Deterrence, "International *Security*, Volume 41, Issue 4, Spring 2017.

[46] August Cole and Amir Husain, "Putin Says Russia's New Weapons Can't Be Beat. With AI and Robotics, They Can," *Defense One*, March 13, 2018, https://www.defenseone.com/ideas/2018/03/putin-says-russias-new-weapons-cant-be-beat-ai-and-robotics-they-can/146631/.

[47] Dave Johnson, *Russia's Conventional Precision Strike Capabilities, Regional Crises, and Nuclear Thresholds*, Livermore Papers on Global Security, No. 3, February 2018; John Warden, *Limited Nuclear War: The 21st Century Challenge for the United States*, Livermore Papers on Global Security, No.4, July 2018. Available at: cgsr.llnl.gov.

[48] Cathy O'Neil, *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*, (New York: Broadway Books, 2017).

[49] Hillary Sanders, Joshua Saxe, "Garbage In, Garbage Out: How Purportedly Great Machine Language Models Can Be Screwed Up by Bad Data," Proceedings of Blackhat 2017, July 26-27, 2017, Las Vegas, NV.

[50] Jesse Emspak, "How a Machine Learns Prejudice," *Scientific American*, December 29, 2016, https://www.scientificamerican.com/article/how-a-machine-learns-prejudice/.

[51] ProPublica, Machine Bias, May 23, 2016, https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing; Will Knight, "Forget Killer Robots, Bias Is the Real AI Danger," *Technology Review*, October 3, 2017, https://www.technologyreview.com/s/608986/forget-killer-robotsbias-is-the-real-ai-danger/.

[52] Louise Matsakis, "Researchers Fooled a Google AI Into Thinking a Rifle Was A Helicopter," Wired, December 20, 2017, https://www.wired.com/story/researcher-fooled-a-google-ai-into-thinking-a-rifle-was-a-helicopter/.

[53] Daniel Cebul, "Differentiating a port from a shipyard is a new kind of problem for AI," C4ISRNET, September 18, 2018, https://www.c4isrnet.com/intel-geoint/2018/09/18/differentiating-a-port-from-a-shipyard-is-a-new-kind-of-problem-for-ai/?utm_source=Sailthru&utm_medium=email&utm_campaign=Daily%209/19&utm_term=Editorial%20-%20Daily%20Brief.

[54] Anna Rohrbach, Lisa Anne Hendricks, Kaylee Burns, Trevor Darrell, Kate Saenko, "Object Hallucination in Image Captioning," Cornell University Library, https://arxiv.org/abs/1809.02156.

[55] Sandia National Laboratory, *Counter Adversarial Data Analytics*, SAND2015-3711, May 8, 2015.

[56] Defense Science Board, Memorandum for Chairman, Terms of Reference, Defense Science Board Task Force on Counter Autonomy, June 18, 2018, https://www.acq.osd.mil/dsb/TORs/2018_TOR_CounterAutonomy_18Jun2018.pdf.

[57] Tim Weiner, "Word for Word, The Cuban Missile Crisis: When Kennedy Faced Armageddon, and His Own Scornful General," *New York Times*, October 5, 1997, https://www.nytimes.com/1997/10/05/weekinreview/word-for-word-cuban-missile-crisis-when-kennedy-faced-armageddon-his-own.html.

[58] Paul Scharre, "A Million Mistakes a Second," *Foreign Policy*, September 12, 2018, https://foreignpolicy.com/2018/09/12/a-million-mistakes-a-second-future-of-war/.

[59] Lawrence Livermore National Laboratory, "Building a Network of Collaborative Autonomous Machines," *Science and Technology Review*, June 2018; Mark Pomerleau, "To win future conflicts, combatant commands must be integrated," C4ISRNET, August 15, 2018, https://www.c4isrnet.com/show-reporter/dodiis/2018/08/14/to-win-future-conflicts-combatant-commands-must-be-integrated/?utm_source=Sailthru&utm_medium=email&utm_campaign=Daily%208/15&utm_term=Editorial%20-%20Daily%20Brief.

[60] Will Knight, "The Dark Secret at the Heart of AI," *Technology Review*, April 11, 2017, https://www.technologyreview.com/s/604087/the-dark-secret-at-the-heart-of-ai/.

[61] Michael Piellusch and Tom Galvin, "Is the Chain of Command Still Meaningful?" War Room, US Army War College, September 6, 2018, https://warroom.armywarcollege.edu/articles/chain-of-command/.

[62] Stimson Center, *An Action Plan on US Drone Policy*: *Recommendations for the Trump Administration*, 2018, https://www.stimson.org/sites/default/files/file-attachments/Stimson%20Action%20Plan%20on%20US%20Drone%20Policy.pdf.

[63] Herb Lin, "Developing Responses to Cyber-Enabled Information Warfare and Influence Operations," Lawfare, September 6, 2018, https://www.lawfareblog.com/developing-responses-cyber-enabled-information-warfare-and-influence-operations.

[64] Kori Schake, "Why We Get It Wrong: Reflecting on the Future of War," book review of Lawrence Freedman, *The Future of War: A History*, *War on the Rocks*, August 10, 2018, https://warontherocks.com/2018/08/why-we-get-it-wrong-reflections-on-predicting-the-future-of-war/; Richard Danzig, *Driving in the Dark: Ten Propositions About Prediction and National Security*, Center for a New American Security, October 2011.

[65] Frank Gac, Timothy Grayson, Joseph Keogh, "What Works? Public-Private Partnerships for Development of National Security Technology," in Davis and Nacht, eds, *Strategic Latency Red, White and Blue*, ibid.

[66] Suzanne Nossel, "Google is Handing the Future of the Internet to China," *Foreign Policy*, September 10, 2018, https://foreignpolicy.com/2018/09/10/google-is-handing-the-future-of-the-internet-to-china/.

[67] Laura Seligman, "Why the Military Must Learn to Love Silicon Valley," September 12, 2018, *Foreign Policy*, https://foreignpolicy.com/2018/09/12/why-the-military-must-learn-to-love-silicon-valley-pentagon-google-amazon/.

[68] Lawrence Freedman, *The Evolution of Nuclear Strategy*, (New York: St. Martin's Press, 1981).

[69] Elsa Kania and John Costello, *Quantum Hegemony" China's Ambitions and the Challenge to US Innovation Leadership*, Center for a New American Security, September 12, 2018, https://www.cnas.org/publications/reports/quantum-hegemony.

[70] Department of Defense, 2018 Nuclear Posture Review, https://dod.defense.gov/News/SpecialReports/2018NuclearPostureReview.aspx.